

基于轻量化 MLCE - RTMDet 的人工去雄后玉米雄穗检测算法

李金瑞^{1,2} 杜建军^{2,3} 张宏鸣¹ 郭新宇^{2,3} 赵春江^{1,3}

(1. 西北农林科技大学信息工程学院, 陕西杨凌 712100; 2. 北京市农林科学院信息技术研究中心, 北京 100097;
3. 国家农业信息化工程技术研究中心, 北京 100097)

摘要: 玉米制种田遗漏雄穗检测是实现人工去雄质量评估的关键。针对现有玉米雄穗检测模型面临的参数量大、检测效率低和精度差等问题,提出一种基于 RTMDet - tiny 的轻量级雄穗检测模型 MLCE - RTMDet。模型采用轻量级的 MobileNetv3 作为主干特征提取网络,有效降低模型参数量;在特征提取网络中引入 CBAM 注意力模块,增强对雄穗目标的多尺度特征提取能力,克服引入轻量化网络可能带来的性能损失。同时,使用 EIOU Loss 替代 GIOU Loss,进一步提高雄穗检测精度。在自建数据集上的试验表明,改进的 MLCE - RTMDet 模型参数量缩减至 3.9×10^6 ,浮点运算数降至 5.3×10^9 ,参数量和浮点运算数分别比原模型减少 20.4% 和 34.6%。测试集上模型平均精度均值增至 92.2%,较原模型提高 1.2 个百分点;同时,推理速度达到 41.9 f/s,增幅达 12.6%。与 YOLO v6、YOLO v8、YOLO X 等当前主流模型相比,MLCE - RTMDet 表现出更好的综合检测性能。改进后的高精度轻量化模型可为实现玉米制种田人工去雄后的遗漏雄穗检测提供技术支撑。

关键词: 无人机; 目标检测; 人工去雄; 玉米雄穗; RTMDet; 轻量化网络

中图分类号: TP391.4 文献标识码: A 文章编号: 1000-1298(2024)11-0184-09

OSID:



Maize Tassel Detection Algorithm after Artificial Emasculation Based on Lightweight MLCE - RTMDet

LI Jinrui^{1,2} DU Jianjun^{2,3} ZHANG Hongming¹ GUO Xinyu^{2,3} ZHAO Chunjiang^{1,3}

(1. College of Information Engineering, Northwest A&F University, Yangling, Shaanxi 712100, China

2. Information Technology Research Center, Beijing Academy of Agriculture and Forestry Sciences, Beijing 100097, China

3. National Engineering Research Center for Information Technology in Agriculture, Beijing 100097, China)

Abstract: Detecting missed tassels is crucial for assessing the quality of artificial emasculation in maize seed production fields. Aiming at the problems of large parameter quantity, low detection efficiency and poor accuracy of the existing maize tassel detection models, a lightweight tassel detection model based on RTMDet - tiny, named MLCE - RTMDet, was proposed. The model used the lightweight MobileNetv3 as the feature extraction network to effectively reduce the model parameters. The CBAM attention module in the neck network was integrated to enhance multi-scale feature extraction capability for tassel objects, overcoming potential performance losses caused by the lightweight networks. Simultaneously, the EIOU Loss was adopted, replacing the GIOU Loss, which further improved the accuracy of tassel detection. Experiments on the self-built dataset showed that the improved MLCE - RTMDet model reduced model parameters to 3.9×10^6 , while the number of floating point operations was lowered to 5.3×10^9 , resulting in a 20.4% reduction in parameters and a 34.6% decrease in computational complexity compared with that of the original model. When evaluated on the test set, the model's mean average precision (mAP) reached 92.2%, reflecting a 1.2 percentage points improvement over the original model. The inference speed was increased to 41.9 frames per second (FPS), representing a 12.6% enhancement. Compared

收稿日期: 2023 - 12 - 27 修回日期: 2024 - 03 - 26

基金项目: 国家重点研发计划项目(2022YFD1900701)、黑龙江省“揭榜挂帅”科技攻关项目(2021XJ05A02)、北京市农林科学院科技创新能力建设专项(KJCX20230429)、国家自然科学基金项目(42377341)和陕西省重点研发计划项目(2023 - YBNY - 217)

作者简介: 李金瑞(1998—),男,硕士生,主要从事深度学习和图像处理研究,E-mail: jerry_li1998@163.com

通信作者: 赵春江(1964—),男,研究员,中国工程院院士,主要从事农业信息化关键技术研究,E-mail: zhaojc@nrcita.org.cn

with current mainstream detection models such as YOLO v6, YOLO v8, and YOLO X, MLCE - RTMDet demonstrated superior overall detection performance. The improved high-accuracy lightweight model offered technical support for tassel re-inspection and emasculation quality assessment in maize seed production fields following artificial emasculation.

Key words: drone; object detection; artificial emasculation; maize tassel; RTMDet; lightweight network

0 引言

玉米是我国的重要食物、饲料和工业原料,在农业生产和国民经济中占据重要地位。目前,我国杂交玉米占总播种面积的 98% 以上。玉米杂交种优势主要表现在杂交种第一代,后代杂交种优势减弱,因此需要年年制种^[1-2]。玉米杂交制种通常选择两个不同品种的玉米作为父母本,父母本按照 1:5 等比例交叉种植。去雄是在抽雄期去除母本雄穗,目前该环节仍主要依赖人工完成^[3]。在玉米制种实践中,单次去雄通常无法清除所有母本雄穗,需要人工多次检查并去除遗漏的雄穗,防止其成熟散粉^[4]。玉米去雄作业对时效性要求高,去雄不及时将造成母本自交,从而降低玉米种子纯度和产量,因此精准高效地检测人工去雄后遗漏的雄穗对评价去雄质量和制定去雄策略具有重要价值。

利用无人机高通量获取制种田冠层图像已成为当前采集大田玉米冠层信息的主要手段。然而,适合在人工去雄后检测遗漏雄穗的模型还尚未报道。在人工去雄的玉米制种田场景中,遗漏的雄穗通常处于抽雄阶段,即雄穗在苞叶中并未抽出,缺少成熟雄穗的颜色和分支特征,这给雄穗精准识别带来挑战。基于传统图像分析技术的方法^[5-7]严重依赖目标颜色、纹理等特征的提取,在检测遗漏雄穗的应用中准确性和鲁棒性难以得到保证。当前,基于深度学习技术对雄穗冠层图像进行识别^[8-9]已成为主流方法,这种基于数据驱动的方法能够更好地捕捉和提取图像中的复杂特征,从而能够更准确地检测雄穗。但这些雄穗检测模型仍存在参数量大和计算量高等问题。雄穗检测模型轻量化将有利于模型在存储和计算资源受限的边缘设备上部署。

实现检测模型轻量化的同时确保模型精度不受影响是一个挑战。针对玉米雄穗检测,已有一些基于轻量级神经网络的检测方法。例如,梁胤豪等^[10]通过将 SSD 算法的特征提取网络替换为 MobileNet,研发了适用于无人机部署的轻量级成熟雄穗检测模型 SSD_MobileNet。张晓勳等^[11]利用 GhostNet 代替 YOLO v4 的特征提取网络,采用 ReLU6 作为激活函数,构建轻量级的成熟雄穗检测模型 G - YOLO。胡阵等^[12]通过采用 PP - LCNet 替换 YOLO v5s 的主干

网络,并结合 SimOTA 的标签匹配策略以及 SIOU Loss,提出轻量级成熟雄穗检测模型 PSOS - YOLO v5s。通常情况下,采用轻量级特征提取网络能够减少模型参数量,但也在一定程度上造成模型检测精度下降。为提高模型精度,一些学者通过引入注意力机制和使用特征融合层等技术来增强模型的特征提取能力。杨蜀秦等^[13]在 CenterNet 的特征提取网络中加入位置编码,在减少模型参数的同时提高检测精度。LIU 等^[14]通过应用双向特征金字塔网络融合不同尺度的特征,并在 YOLO v5 的基础上增加检测头,提升了对小尺寸雄穗的检测精度。SONG 等^[15]在 YOLO X 中嵌入 SE 注意力机制,提升模型对复杂自然环境中雄穗的检测精度。上述优化策略有利于提高玉米雄穗检测精度,但这些模型的检测对象仅限于特征明显的成熟雄穗,并不适用于人工去雄后场景的未成熟雄穗检测。因此,为实现人工去雄遗漏雄穗的高精度和高效率检测,需进一步研究更适用的雄穗检测模型。

本文以制种田人工去雄后场景的雄穗为研究对象,首先根据人工去雄阶段雄穗的不同状态进行分类,然后改进 RTMDet - tiny 的目标检测模型,在实现轻量化的情况下保证模型检测精度,最后进行模型在无人机采集制种田玉米冠层图像中的应用。

1 材料与方法

1.1 数据集构建

于 2022 年 6 月 30 日—7 月 4 日在新疆维吾尔自治区塔城地区开展玉米制种田冠层图像采集工作。使用大疆无人机 (DJI Matrice 300 RTK 型) 作为数据采集设备,该无人机使用全画幅 Zenmuse P1 型相机,原生像素为 4 500 万,可以拍摄分辨率为 8 192 像素 × 5 460 像素的图像。无人机飞行高度设置为 20 m,航向重叠度和旁向重叠度分别设置为 70% 和 80%,飞行速度为 2 m/s,单次飞行可采集玉米冠层图像约 1 000 幅。相机拍摄角度保持垂直向下,图像地面采样距离为 0.25 cm/像素。

从收集到的图像中筛选出人工去雄后场景的图像共 600 幅。由于无人机影像尺寸较大,无法直接用于模型的训练和检测,因此从每幅无人机图像中

心区域裁剪出尺寸为 1 024 像素 × 1 024 像素的图像构建数据集,用于雄穗标注。通过观察去雄时期玉米雄穗的特点,将雄穗状态分为 Tassel_N 和 Tassel_S 两类,如图 1 所示。其中,Tassel_N 表示玉米雄穗已被去除,Tassel_S 表示玉米雄穗未被去除。



图 1 人工去雄后制种田间两种不同状态雄穗

Fig.1 Tassels in two different states in seed production field after artificial emasculation

1.2 数据增强

自然环境中采集的制种玉米冠层图像易受光照等不利因素的干扰,因此采用数据增强技术^[16]来生成更多有价值的训练样本,以提高雄穗检测模型的稳定性与泛化性。本文使用随机旋转、加入阴影等多种数据增强技术处理训练数据集。如图 2 所示,这些增强策略一方面有助于减弱由于样本规模等问

题导致的模型参数提取失真或过度拟合;另一方面,通过加入阴影、调整图像亮度及对比度来模拟复杂的噪声干扰和光照环境,可以有效解决图像噪点、模糊等质量问题,从而提高训练模型抗干扰能力。本文对训练集中每幅图像随机应用数据增强技术,共生成 480 幅增强后的图像。最终,用于雄穗检测模型训练、验证和测试图像分别为 960、60、60 幅。

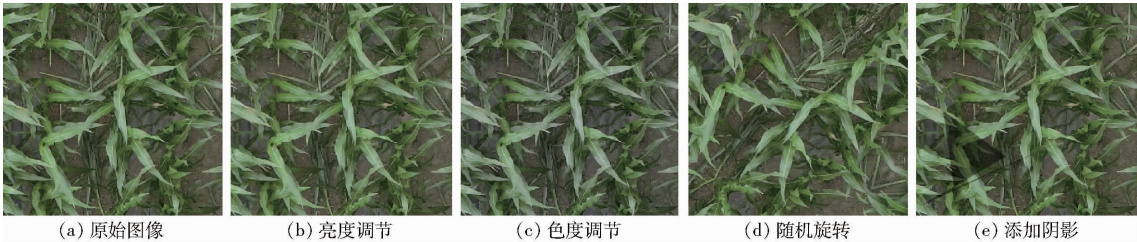


图 2 数据增强后图像

Fig.2 Images after data augmentation

1.3 数据集分析

数据集中所有雄穗的标注框信息统计如表 1 所示。在雄穗类别分布上,Tassel_N 占比 58.6%,Tassel_S 占比 41.4%。所有标注框平均面积和边长分别为 6 966 像素和 83 像素,最小的标注框面积约为 1 101 像素。根据 COCO 数据标注规范^[17],可将尺寸小于 32 像素 × 32 像素的雄穗界定为小目标,

尺寸大于 96 像素 × 96 像素的雄穗视为大目标,其余则归为中等目标。可以看出,本文标注的雄穗目标均属于大/中等目标,不存在小目标。因此,采用 MobileNetv3 作为特征提取网络提取出雄穗多尺度特征。即使 MobileNetv3 的最后一层将特征图压缩至原图的 1/32,依然能够从每个雄穗目标提取出有效信息,从而保证雄穗多尺度特征提取

表 1 数据集标注框统计分析

Tab.1 Statistical analysis on annotation boxes of datasets

类别	训练集标注数	验证集标注数	测试集标注数	总标注数	标注框面积/像素		
					均值	最小值	标准差
Tassel_N	24 255	1 486	1 124	26 865	6 496	1 101	1 427
Tassel_S	17 527	1 055	415	18 997	7 731	1 368	1 680
总体目标	41 782	2 541	1 539	45 862	6 966	1 101	2 276

的完整性。

1.4 RTMDet - tiny 网络结构

RTMDet^[18]是一种基于 YOLO 系列模型进行优化的实时物体检测器,在物体检测中展现出优越的

检测性能,并具备向其他对象识别任务拓展的潜能,如实例分割及旋转物体检测等。选择 RTMDet - tiny 作为基准模型,该模型是 RTMDet 系列的最小化版本,具有体量小巧、参数精简、对硬件要求相对较低

等优点,并且在同参数规模的检测模型中,具有较高的检测精度。RTMDet - tiny 模型的网络架构主要由特征提取网络 (Backbone 层)、颈部网络 (Neck 层) 以及检测头部 (Head 层) 3 部分构成。首先由多层 Conv、CSP 和 SPPF 模块构成的 Backbone 层提取出模型输入图像的深层特征。然后,这些深层特征所

生成的特征图注入 Neck 层分支,该分支继承特征金字塔网络 (Feature pyramid network, FPN) 和路径聚合网络 (Path aggregation network, PAN) 的特点,有效地实现了多尺度特征的融合。最终,经融合处理后的特征图被传送至 Head 层预测模块,进行物体分类和定位。RTMDet 的网络结构如图 3 所示。

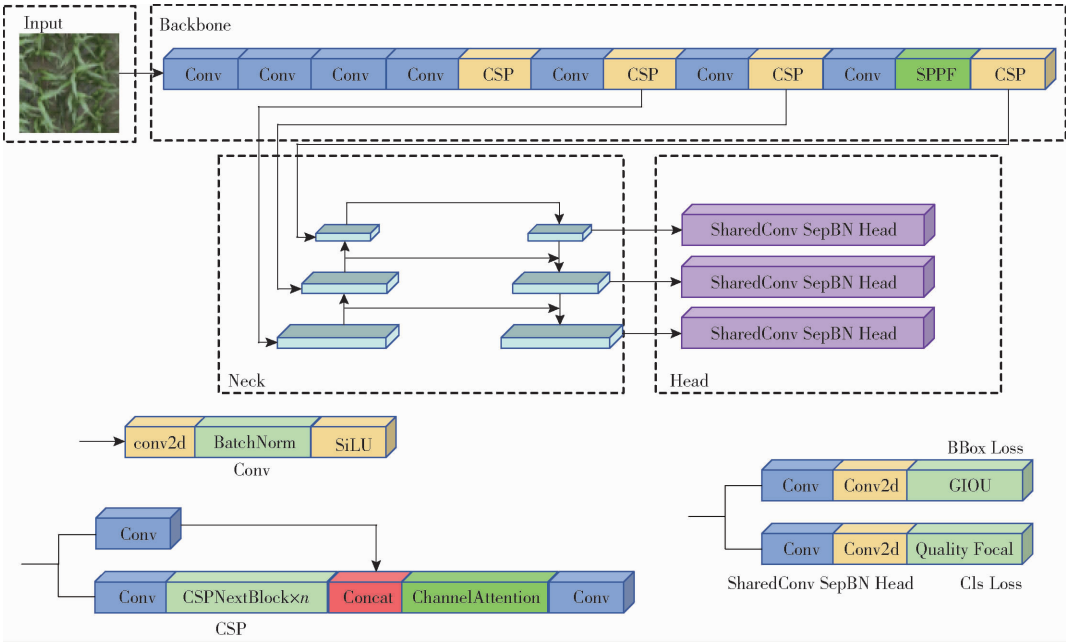


图 3 RTMDet 网络结构图
Fig. 3 Network structure of RTMDet

1.5 RTMDet - tiny 网络改进

为提高人工去雄后遗漏雄穗的模型检测效率,同时保证雄穗检测精度,本文提出一种轻量级网络结构 MLCE - RTMDet,如图 4 所示。改进模型利用 RTMDet - tiny 的网络结构,在主干网络中采用轻量级 MobileNetv3 提取特征,减少模型参数量,提升模

型检测速度。由于轻量化操作可能导致模型检测精度下降,因此在特征提取网络的末端引入注意力机制 CBAM,增强网络提取多尺度特征的能力。在模型 Head 层使用 EIOU Loss 替换 GIOU Loss,进一步提高模型检测精度。这些改进使得提出的模型能够同时保证检测精度和效率。

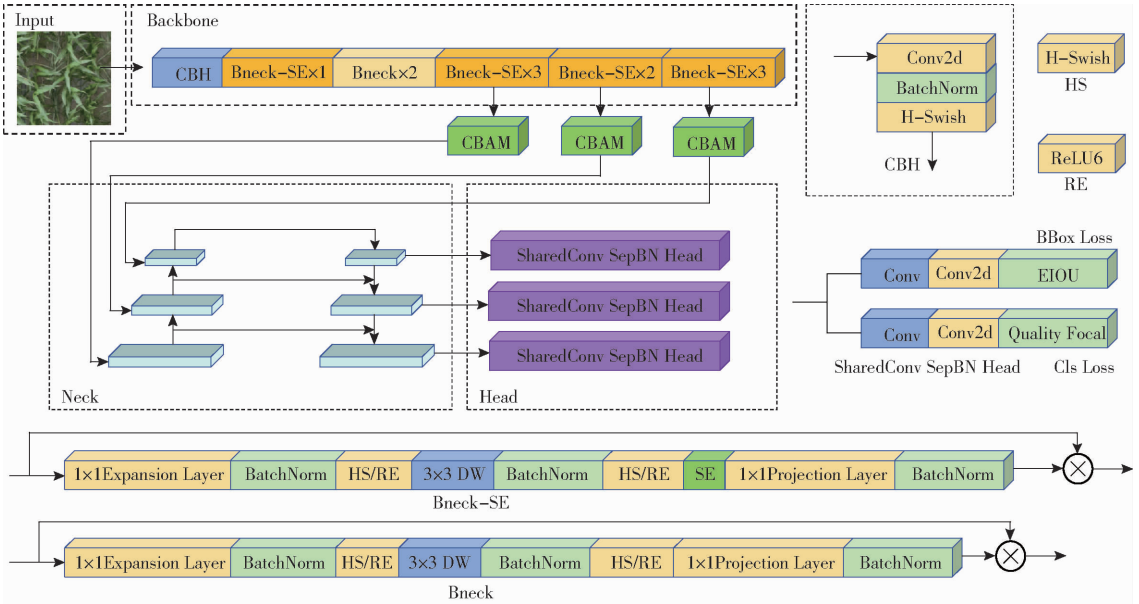


图 4 MLCE - RTMDet 网络结构图
Fig. 4 Network structure of MLCE - RTMDet

1.5.1 网络轻量化

RTMDet-tiny 的主干特征提取网络 CSPNext 中使用了大核深度卷积来提高模型捕获全局上下文的能力,这种方法会增加模型的参数量。为减少网络结构的整体参数量,本文使用轻量化特征提取网络 MobileNetv3 替换 CSPNext。

MobileNet 是一种经典的轻量级特征提取网络结构,已经推出 3 个版本: MobileNetv1、MobileNetv2 和 MobileNetv3。MobileNetv1^[19] 采用深度可分离卷积减少参数量和计算复杂度。MobileNetv2^[20] 使用线性瓶颈、倒残差结构和 ReLU6 激活函数优化特征提取。MobileNetv3^[21] 结合前两版特点,进

一步引入神经架构搜索技术 (Neural architecture search, NAS) 来构建网络,并添加 SE 注意力机制和 H-Swish^[22] 激活函数以提高性能和降低计算成本。

MobileNet3 分为 large 和 small 两个版本,其整体结构一致,但基本单元 Bneck 和 Bneck-SE 的数量及其内部参数存在差异,如图 5 所示。其中, Bneck-SE 作为 MobileNet3 的核心部分,主要由通道可分离卷积、SE 注意力机制以及残差连接组成。选用 MobileNet3-large 替换 RTMDet-tiny 的特征提取网络以实现模型轻量化,图 5 为 MobileNet3 网络结构。

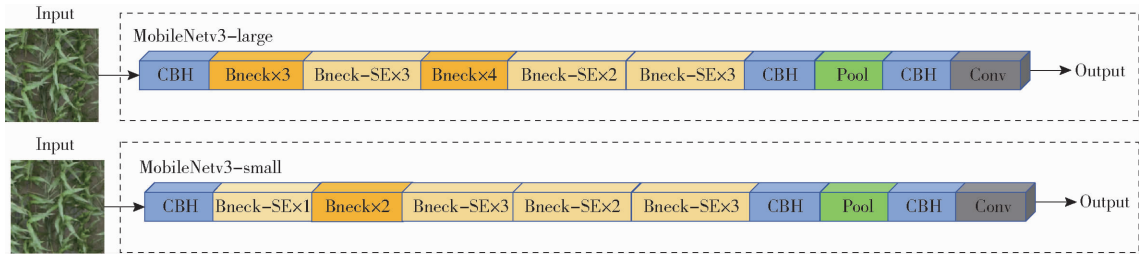


图 5 MobileNetv3 网络结构
Fig.5 Structure of MobileNetv3

1.5.2 增加注意力机制

玉米种植过程中,由于植株密度过高导致叶片间相互遮挡,这使得轻量化骨干网络在准确提取关键信息时面临挑战。为解决这一难题,本文引入注意力机制,更有效地聚焦和提取图像中的关键信息。注意力机制能够对图像不同区域进行加权处理,从而更精确地捕获重要特征。CBAM^[23] (Convolutional block attention module) 是一种具有代表性的注意力机制,它在跨维度注意力权重与输入特征的整合方面表现出巨大潜力。该机制结合卷积操作与注意力机制的特点,能够从空间和通道两个维度对图像进行关注,并自适应地学习输入特征图的重要性,以增强特征表达能力。CBAM 模块设计相对简洁,所增加的计算量和参数量相对较少,这使得它可以轻松地集成到任何卷积神经网络架构中。因此,CBAM 是一种应用广泛且高效的注意力机制,如图 6 所示。

1.5.3 损失函数调整

头部网络的损失函数直接关系到检测准确率。RTMDet 中的损失函数由分类损失函数 Quality FocalLoss^[24] 和边界框损失函数 GIOU Loss^[25] 两部分组成。

根据表 1 中的目标数量统计结果,Tassel_N 的数量高于 Tassel_S,存在目标比例失衡问题,使用分类损失函数 Quality FocalLoss 能够较好地解决该问题,因此本文保留 RTMDet 算法中的该分类损失

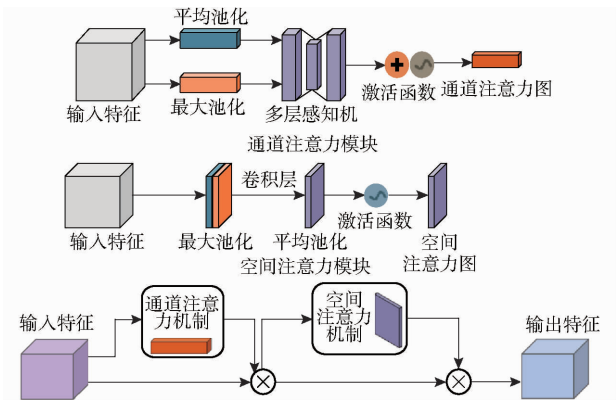


图 6 CBAM 结构图
Fig.6 Structure of CBAM

函数。

在边界框损失的计算中,选择合适的方法对于提升目标检测模型的性能非常重要。一个好的边界框损失函数应该考虑预测框和真实框的重叠度^[26]、中心点距离^[27] 和长宽比的一致性^[28]。RTMDet 模型使用的边界框损失函数是 GIOU Loss,它是基于距离度量的损失函数,对尺度并不敏感。尤其是当检测框和真实框存在包含关系时,GIOU 无法准确区分检测框和真实框的相对位置关系,从而使优化变得困难。因此,本文采用包含重叠损失、中心距离损失和宽高损失 3 部分的 EIOU Loss 替代 GIOU Loss。这样的改进可以更全面地评价预测框和真实框之间的关系,进而提升目标检测模型性能。

2 试验与结果分析

2.1 试验环境

硬件环境为搭载 Intel (R) Core (TM) i7 - 10700CPU@2.90 GHz 和 Nvidia GeForce RTX 1660s GPU 的计算机。操作系统为 Windows 10,在安装 CUDA 11.1 和 cuDNN 8.0.5 的 Python 环境下完成深度学习模型的开发、训练和推理。模型训练的超参数设置如下:批处理大小设置为 4,初始学习率为 4×10^{-3} ,权重衰减为 0.05,AdamW 动量为 0.9,非最大抑制阈值为 0.6,训练迭代次数为 300。

2.2 评价指标

选取平均精度均值 (Mean average precision, mAP)、准确率 (Accuracy)、参数量 (Parameters)、浮点运算数 (Floating point operations, FLOPs) 和每秒处理帧数 (Frames per second, FPS) 作为评价指标评估雄穗检测模型性能。

2.3 消融试验

在相同超参数设置下,以 RTMDet - tiny 为基础模型 (mAP 为 91%),开展消融试验:轻量化骨干网络消融 (将骨干网络替换为 MobileNetv3)、注意力机制消融 (引入注意力机制 CBAM)、损失函数消融 (将头部网络的损失函数从 GIOU 更改为 EIOU) 和雄穗检测模型消融,结果如表 2 ~ 5 所示。

表 2 轻量化骨干网络消融试验结果

Tab.2 Results of ablation experiments for lightweight backbone network			
模型	参数量	浮点运算数	mAP@0.5/%
RTMDet - tiny	4.9×10^6	8.1×10^9	91.0
+ M3l	3.9×10^6	5.3×10^9	90.9
+ M3s	1.6×10^6	3.7×10^9	88.9

表 3 注意力机制消融试验结果

Tab.3 Results of ablation experiments for attention mechanism			
模型	参数量	浮点运算数	mAP@0.5/%
RTMDet - tiny	4.9×10^6	8.1×10^9	91.0
+ CBAM	4.9×10^6	8.1×10^9	91.6
+ M3s + CBAM	1.6×10^6	3.7×10^9	89.7
+ M3l + CBAM	3.9×10^6	5.3×10^9	91.1

2.3.1 轻量化骨干网络消融试验

根据上文所述, MobileNet3 有两个不同规模的版本,将主干网络替换为 MobileNet3 - large,标记为 + M3l;将模型的主干网络替换为 MobileNet3 - small,标记为 + M3s。将特征提取网络进行调整后,模型性能如表 2 所示。将 RTMDet - tiny 的特征提取网络替换为 MobileNet3 - large 后,模型参

数量减少 20.4%,浮点运算数减少 34.6%。在测试集上, mAP@0.5 降至 90.9%, mAP@0.5 下降幅度较小,表明骨干网络轻量化对性能影响较小。然而,当将 RTMDet - tiny 的特征提取网络替换为 MobileNetv3 - small 后,模型总体参数量显著减少 67.3%,模型 mAP@0.5 也显著下降,降至 88.9%。使用轻量化特征提取网络,虽然有效地降低了模型参数量,但也导致模型 mAP@0.5 不同程度地下降。

表 4 损失函数消融试验结果

Tab.4 Results of ablation experiments for loss function

模型	参数量	浮点运算数	mAP@0.5/%
RTMDet - tiny	4.9×10^6	8.1×10^9	91.0
+ EIOU	4.9×10^6	8.1×10^9	91.7
+ M3s + EIOU	1.6×10^6	3.7×10^9	89.7
+ M3l + EIOU	3.9×10^6	5.3×10^9	92.0

表 5 雄穗检测模型消融试验结果

Tab.5 Results of ablation experiments for tassels detection model

模型	参数量	浮点运算数	mAP@0.5/%
RTMDet - tiny	4.9×10^6	8.1×10^9	91.0
+ M3s + CBAM + EIOU	1.6×10^6	3.7×10^9	90.3
+ M3l + CBAM + EIOU	3.9×10^6	5.3×10^9	92.2

2.3.2 注意力机制消融试验

为弥补由模型轻量化导致的平均精度均值下降,本文通过在特征提取网络末端引入注意力机制 (CBAM),充分挖掘骨干特征输出的有效信息。模型添加 CBAM 后的检测性能如表 3 所示。在基准模型 RTMDet - tiny 中添加 CBAM 后, mAP@0.5 从 91% 提至 91.6%。在更换特征提取网络模型中添加 CBAM 后, RTMDet - tiny + M3l 的 mAP@0.5 从 90.9% 提至 91.1%, RTMDet - tiny + M3s 的 mAP@0.5 从 88.9% 提至 89.7%。添加注意力机制 CBAM 后,模型 mAP@0.5 得到进一步提高,模型参数量未明显增加,保持了模型轻量化特性。为了直观地显示在骨干网络特征输出末端引入 CBAM 注意力机制对雄穗检测模型检测性能的影响,使用热力图可视化模型的注意力区域^[29],如图 7 所示。图中,颜色类别表示是否关注目标 (红色为关注,蓝色为不关注),颜色深浅反映关注程度 (颜色越深表示程度越高)。比较改进后模型和基准模型热力图,可看出基础模型关注过多的背景区域,改进后模型更关注玉米株心目标。因此,通过引入注意力机制,能够使轻量化模型实现更好的检测性能。

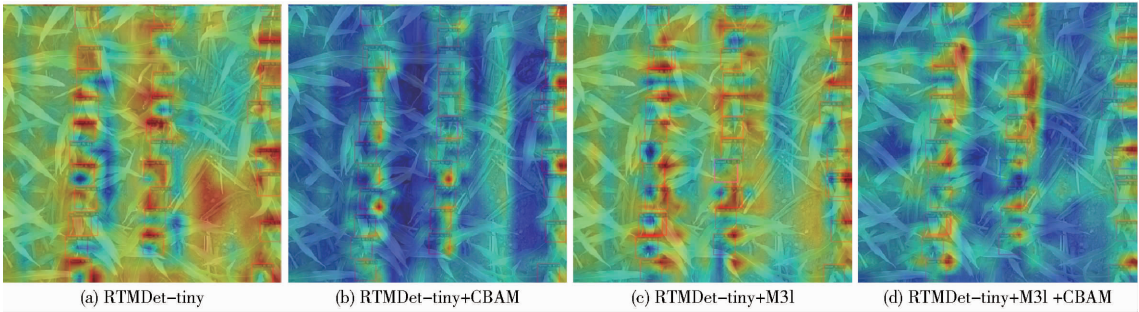


图 7 注意力机制消融试验热图

Fig. 7 Heat maps resulted from ablation experiments for attention mechanism

2.3.3 损失函数消融试验

损失函数改进体现在检测框损失上。由于 EIOU 损失函数更全面地考虑了不同情况下的损失,将原头部结构的 GIOU 损失函数替换为 EIOU 损失函数后,模型 mAP@0.5 从 91% 提至 91.7%,如表 4 所示。在更换特征提取网络后的轻量化模型上修改边界框损失函数后,RTMDet-tiny+M3l mAP@0.5 从 90.9% 提升至 92%,RTMDet-tiny+M3s mAP@0.5 从 88.9% 提升至 89.7%。可以看出,更换损失函数后,所有模型的平均精度均值都得到提升,但参数量未发生显著变化。

2.3.4 雄穗检测模型消融试验

从上述试验可知,嵌入注意力机制和修改损失函数对模型参数量和浮点运算数影响甚微,这些改

进显著提升模型的检测性能。这不仅验证了所提出改进措施的有效性,也为进一步优化目标检测模型提供了有价值的参考。由表 5 可知,改进后的算法 MLCE-RTMDet (RTMDet-tiny+M3l+CBAM+EIOU) 的 mAP@0.5 提高 1.2 个百分点,整体参数量减少 20.4%。MSCE-RTMDet (RTMDet-tiny+M3s+CBAM+EIOU) 整体参数量下降 67.3%,模型 mAP@0.5 也有所下降。这表明在模型轻量化的同时,可能损失部分模型精度。因此,模型的轻量化和高精度之间存在一定的权衡关系,需要在实际应用中进行平衡和选择。MLCE-RTMDet 既提升了模型的检测性能,又达到轻量化算法的设计目标。图 8 为改进后模型热力图,直观地显示了模型

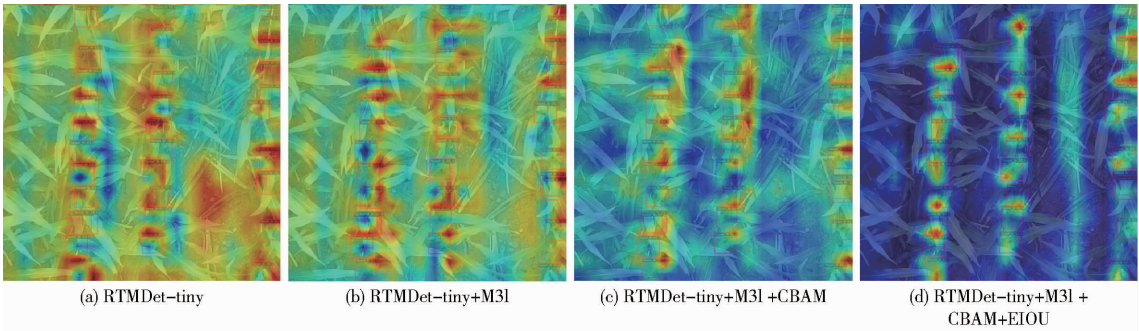


图 8 雄穗检测模型消融试验热图

Fig. 8 Heat maps resulted from ablation experiments for tassel detection model

对目标关注程度。对比不同模型热力图,可以看出各项改进均使模型更加注重雄穗目标的特征,并能有效抑制对叶片和地面等背景区域的关注。

因此,本文提出的模型优化方法对人工去雄后遗漏的雄穗检测具有价值。

2.4 雄穗检测模型性能对比

为评估改进后 MLCE-RTMDet 模型在玉米雄穗检测任务中的综合性能,选取 YOLO v6、YOLO v8、YOLO X 和 Faster R-CNN 等主流目标检测模型在测试集上进行对比试验,结果如表 6 所示。可以看出,模型 MLCE-RTMDet 检测精度 (mAP@0.5) 指标优于其它模型,且运算数和参数量指标保持在较低水平。与基准模型 RTMDet-tiny 的推理速度

相比,模型 MLCE-RTMDet 的推理速度提升 12.6%。

表 6 雄穗检测模型性能对比

Tab. 6 Performance comparison on different tassel detection models

模型	参数量	浮点运算数	推理速度/ mAP@0.5/	
			(f·s ⁻¹)	%
YOLO v6	4.30 × 10 ⁶	5.50 × 10 ⁹	35.5	87.4
YOLO v8	3.10 × 10 ⁶	4.10 × 10 ⁹	75.0	88.4
YOLO X	1.40 × 10 ⁶	1.90 × 10 ⁹	57.3	88.5
Faster R-CNN	4.12 × 10 ⁷	7.82 × 10 ¹⁰	16.7	85.1
RTMDet-tiny	4.90 × 10 ⁶	8.10 × 10 ⁹	37.2	91.0
MSCE-RTMDet	1.60 × 10 ⁶	3.70 × 10 ⁹	43.0	90.3
MLCE-RTMDet	3.90 × 10 ⁶	5.30 × 10 ⁹	41.9	92.2

为直观验证改进模型的检测效果,对比了 MLCE - RTMDet 与 RTMDet - tiny 模型在测试图像上的检测结果,如图 9 所示。图中,红色检测框表示 Tassel_N 类目标,紫色检测框表示 Tassel_S 类目标;红色箭头所指为漏检目标,第 1 行箭头指示 Tassel_N 类型漏检目标,其它行箭头指示 Tassel_S 类型漏检

目标。检测结果显示,尽管 RTMDet - tiny 模型已经具有较高的检测精度,但对 Tassel_N 和 Tassel_S 两类目标仍存在明显漏检现象(图 9b)。相较而言,MLCE - RTMDet 在同样的测试图像上表现出更好的检测性能,能够准确识别出 RTMDet - tiny 模型遗漏的目标(图 9c)。

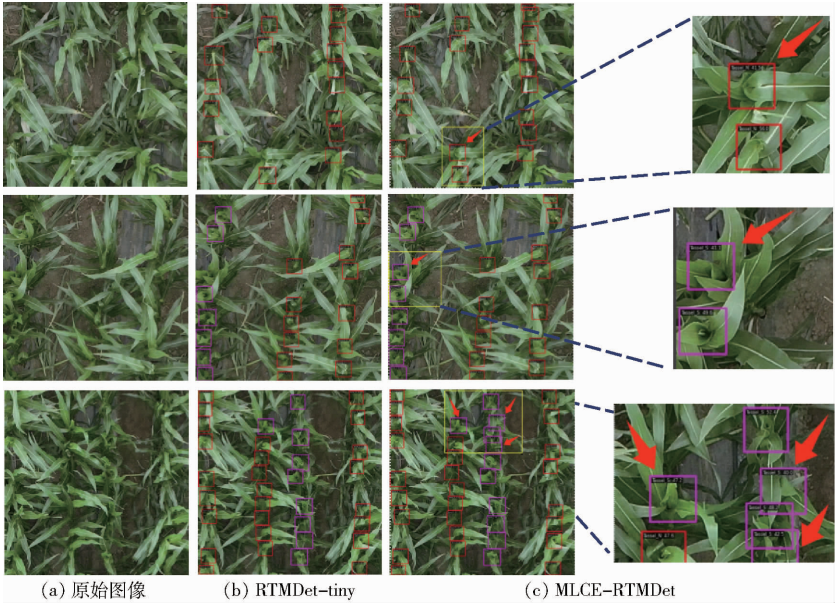


图 9 雄穗检测模型在测试集上检测效果对比

Fig. 9 Comparison of detection results of tassel detection models on test set

2.5 雄穗检测模型应用

由于制种田获取的无人机图像尺寸远超模型推理适用的输入图像尺寸,需要将其切分成小图像送入雄穗检测模型。无人机图像远离中心位置的分块图像上,雄穗目标存在明显畸变,使得检测模型在这些分块图像上检测准确率差异较大(图 10)。以各个图像分块的雄穗检测准确率作为评价指标,评估 MLCE - RTMDet 模型在无人机采集的原始图像上的雄穗检测性能及其适用范围。

为保持与雄穗检测模型的训练图像尺寸一致,将无人机原始图像裁剪为 40 个尺寸为 1 024 像素 × 1 024 像素的图像分块,并按照行列号进行相应标记,如图 10 所示。统计雄穗检测模型各分块图像检测准确率,如表 7 所示。可以看出,模型对无人机图像中心附近的 3 × 4 个分块图像的平均检测准确率超过 90%。其它位置分块图像,除位于第 1 列和第 8 列的分块外,模型准确率也超过 80%。通常情况下,距离图像中心点越近的分块,雄穗检测准确率越高。进一步比较不同位置的图像分块(分块图像 2 - 1 和 3 - 5),可以看出:边缘分块 2 - 1 区域中玉米植株更容易被叶片遮挡,导致模型检测准确率下降。为将模型推向应用,需要在无人机采集图像时设置合适的重叠率,以保证模型在每块区域上的

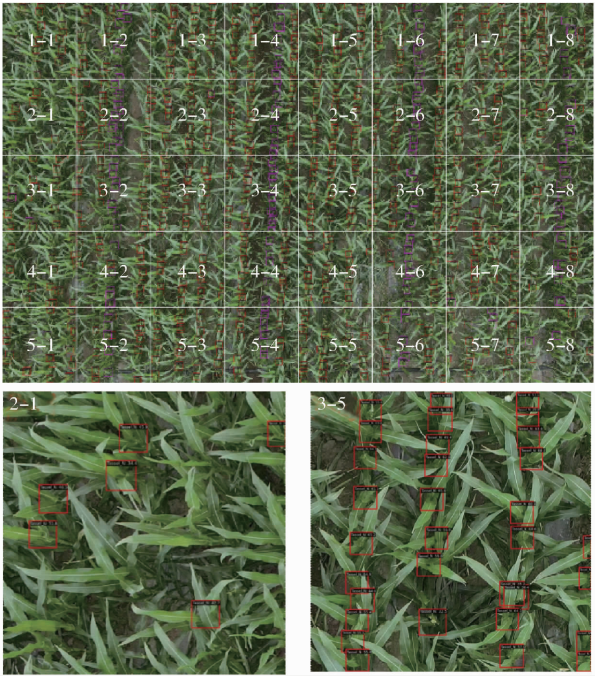


图 10 雄穗检测模型在无人机图像上的应用

Fig. 10 Application of tassel detection model on UAV image

检测准确率。

3 结论

(1) 针对现有玉米雄穗检测模型面临的参数量

表 7 无人机图像分块上雄穗检测准确率统计结果

Tab. 7 Detection accuracy statistics for UAV								
image blocks								%
行号	列号							
	Col-1	Col-2	Col-3	Col-4	Col-5	Col-6	Col-7	Col-8
Row-1	67	81	88	95	93	94	92	74
Row-2	25	97	97	100	97	93	87	54
Row-3	78	94	97	100	100	100	93	64
Row-4	65	81	89	96	96	95	93	47
Row-5	67	82	93	92	97	85	84	25

注:Col-1 表示第 1 列,Row-1 表示第 1 行,以此类推。

大、检测效率低和精度差等问题,提出一种轻量级的雄穗检测模型。模型参数数量为 3.9×10^6 ,模型浮点运算数为 5.3×10^9 ,推理速度为 41.9 f/s,可应用于人工去雄后遗漏雄穗的精准检测。

(2)将雄穗检测范围从成熟雄穗扩展到人工去雄时期的雄穗,将雄穗状态分为已去除雄穗和未去除雄穗两类,构建制种田玉米去雄阶段的雄穗检测

数据集。

(3)使用 MobileNetv3-large 轻量化网络替换了 RTMDet-tiny 的特征提取网络 CSPNext,轻量化后的模型相比基准模型,参数量减少 20.4%,浮点运算数减少 34.6%。为消除轻量化后对雄穗检测模型检测精度的影响,将 CBAM 注意力机制嵌入特征提取网络末端,使模型在学习时更加关注雄穗特征。在 RTMDet-tiny 的检测头中,采用 EIOU 损失函数替换原有的 GIOU 损失函数,进一步提高模型检测精度。优化后轻量化模型 MLCE-RTMDet 在测试集上对雄穗的平均精度均值可达 92.2%,优于基准模型与目前主流的检测模型,且在轻量化方面具有优势。

(4)针对实际应用需求,通过对无人机原始图像进行分块处理,并对不同位置分块图像进行了检测准确率对比,发现成像畸变对边缘位置分块图像的雄穗检测具有极为不利的影响。因此需要制定合适的重叠率以确保雄穗检测模型应用效果。

参 考 文 献

[1] RANUM P, PEÑA-ROSAS J, GARCIA-CASAL M. Global maize production, utilization, and consumption[J]. Annals of the New York Academy of Science, 2014,1312(1):105-112.

[2] ERENSTEIN O, JALETA M, SONDER K, et al. Global maize production, consumption and trade: trends and R&D implications[J]. Food Security,2022,14(5):1295-1319.

[3] CZEPAK M, KLIEMANN M, SCHMILDT O, et al. Effect of artificial detasseling and defoliation on maize seed production[J]. International Journal of Plant & Soil Science, 2019, 20:1-9.

[4] 丁帅涛, 孙琴, 罗红兵. 玉米雄穗分化发育研究进展[J]. 作物研究, 2014, 28(1): 97-102.
DING Shuaitao, SUN Qin, LUO Hongbing. Research progress on differentiation and development of tassel in maize[J]. Crop Research, 2014, 28(1): 97-102. (in Chinese)

[5] TANG W, ZHANG Y, ZHANG D, et al. Corn tassel detection based on image processing[C]//2012 International Workshop on Image Processing and Optical Engineering. SPIE, 2012:123-129.

[6] YE M, CAO Z, YU Z, et al. An image-based approach for automatic detecting tasseling stage of maize using spatio-temporal saliency[C]//Remote Sensing Image Processing, Geographic Information Systems, and other Applications, 2013: 8921.

[7] KURTULMUŞ F, KAVDIR İ. Detecting corn tassels using computer vision and support vector machines[J]. Expert Systems with Applications, 2014, 41(16): 7390-7397.

[8] 马中杰, 罗晨, 骆巍, 等. 基于改进 YOLO v7-tiny 的玉米种质资源雄穗检测方法[J]. 农业机械学报,2024,55(7):290-297.
MA Zhongjie, LUO Chen, LUO Wei, et al. Tassel detection method of maize germplasm resources based on improved YOLO v7-tiny[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024,55(7):290-297. (in Chinese)

[9] 常婉秋, 姚宇, 席晓杰, 等. 基于迁移学习和非监督分类的制种玉米遥感识别方法[J]. 农业机械学报,2024,55(8):181-195.
CHANG Wanqiu, YAO Yu, XI Xiaojie, et al. Remote sensing identification for seed maize with integrated migration learning and unsupervised classification[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024,55(8):181-195. (in Chinese)

[10] 梁胤豪, 陈全, 董彩霞, 等. 基于深度学习和无人机遥感技术的玉米雄穗检测研究[J]. 福建农业学报, 2020,35(4): 456-464.
LIANG Yinhao, CHEN Quan, DONG Caixia, et al. Application of deep-learning and UAV for field surveying corn tassel[J]. Fujian Journal of Agricultural Sciences, 2020, 35(4): 456-464. (in Chinese)

[11] 张晓勐, 朱德利, 余茂生. 无人机遥感图像中的玉米雄穗轻量化检测模型[J]. 江西农业大学学报,2022,44(2):461-472.
ZHANG Xiaomeng, ZHU Deli, YU Maosheng. Lightweight detection model of maize tassel in UAV remote sensing image[J]. Acta Agriculturae Universitatis Jiangxiensis, 2022, 44(2): 461-472. (in Chinese)

[12] 胡阵, 马宗军, 黄传宝, 等. PSOS-YOLO v5s: 一种轻量级玉米雄穗快速检测算法[J/OL]. 无线电工程. <https://link.cnki.net/urlid/13.1097.TN.20230915.1739.002>.
HU Zhen, MA Zongjun, HUANG Chuanbao, et al. PSOS-YOLO v5s: a lightweight maize tassel rapid detection algorithm [J/OL]. Radio Engineering. <https://link.cnki.net/urlid/13.1097.TN.20230915.1739.002>. (in Chinese)

[13] 杨蜀秦, 刘江川, 徐可可, 等. 基于改进 CenterNet 的玉米雄蕊无人机遥感图像识别[J]. 农业机械学报,2021,52(9):206-212.
YANG Shuqin, LIU Jiangchuan, XU Keke, et al. Improved CenterNet based maize tassel recognition for UAV remote sensing image[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(9): 206-212. (in Chinese)

[18] 张兵, 黄华, 蔡佳敏, 等. 冗余六自由度并联隔振平台多维隔振性能研究[J]. 振动与冲击, 2022, 41(2): 26–32.
ZHANG Bing, HUANG Hua, CAI Jiamin, et al. Multi-dimensional isolation performances of a redundant 6DOF parallel vibration isolation platform[J]. Journal of Vibration and Shock, 2022, 41(2): 26–32. (in Chinese)

[19] 张英, 孙豪, 马帅帅. 基于 4-UPS/CPC 并联机构的多维调姿隔振平台设计[J]. 中国机械工程, 2021, 32(13): 1513–1522, 1529.
ZHANG Ying, SUN Hao, MA Shuaishuai. Design of a multi-dimensional attitude adjustment and vibration isolation platform based on 4-UPS/CPC parallel mechanism[J]. China Mechanical Engineering, 2021, 32(13): 1513–1522, 1529. (in Chinese)

[20] 耿明超, 赵铁石, 李二伟, 等. 主被动闭环融合调姿隔振平台动力学建模与模态分析[J]. 农业机械学报, 2022, 53(7): 435–446.
GENG Mingchao, ZHAO Tieshi, LI Erwei, et al. Dynamic modeling and modal analysis of orientation adjustment and vibration isolation platform with active-passive closed-loop fusion[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(7): 435–446. (in Chinese)

[21] 黄真, 赵永生, 赵铁石. 高等空间机构学[M]. 北京: 高等教育出版社, 2006: 73–83.

[22] 张雷雷, 赵延治, 赵铁石. 并联机构瞬轴面研究进展[J]. 机械工程学报, 2023, 59(21): 131–146.
ZHANG Leilei, ZHAO Yanzhi, ZHAO Tieshi. State of the art of axodes traced by parallel mechanism[J]. Journal of Mechanical Engineering, 2023, 59(21): 131–146. (in Chinese)

[23] ZHAO Tieshi, GENG Mingchao, CHEN Yuhang, et al. Kinematics and dynamics hessian matrices of manipulators based on screw theory[J]. Chinese Journal of Mechanical Engineering, 2015, 28(2): 226–235.

[24] 赵星宇, 赵铁石, 徐雪寒, 等. 3-P(4S) 并联平台振动特性分析与实验验证[J]. 农业机械学报, 2018, 49(7): 419–426.
ZHAO Xingyu, ZHAO Tieshi, XU Xuehan, et al. Vibration characteristics analysis and experimental validation of 3-P(4S) parallel platform[J]. Transactions of the Chinese Society for Agricultural Machinery, 2018, 49(7): 419–426. (in Chinese)

[25] RAO S S. 机械振动[M]. 5 版. 李欣业, 杨理诚, 译. 北京: 清华大学出版社, 2016: 208–210.

(上接第 192 页)

[14] LIU W, QUIJANO K, CRAWFORD M. YOLO v5 – Tassel: detecting tassels in RGB UAV imagery with improved YOLO v5 based on transfer learning[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2022, 15: 8085–8094.

[15] SONG C, ZHANG F, LI J, et al. Detection of maize tassels for UAV remote sensing image with an improved YOLO X model [J]. Journal of Integrative Agriculture, 2023, 22(6): 1671–1683.

[16] SHORTEN C, KHOSHGOFTAAR T. A survey on image data augmentation for deep learning[J]. Journal of Big Data, 2019, 6(1): 60–108.

[17] LIN T Y, MAIRE M, BELONGIE S, et al. Microsoft COCO: common objects in context [C] // Proceedings of European Conference on Computer Vision, 2014: 740–755.

[18] LYU C, ZHANG W, HUANG H, et al. RTMDet: an empirical study of designing real-time object detectors [J]. arXiv Preprint, arXiv: 2212.07784, 2022.

[19] HOWARD A G, ZHU M, CHEN B, et al. MobileNets: efficient convolutional neural networks for mobile vision applications [J]. arXiv Preprint, arXiv: 1704.04861, 2017.

[20] SANDLER M, HOWARD A, ZHU M, et al. MobileNetV2: inverted residuals and linear bottlenecks [J]. arXiv Preprint, arXiv: 1801.04381, 2019.

[21] HOWARD A, SANDLER M, CHU G, et al. Searching for MobileNetV3 [J]. arXiv Preprint, arXiv: 1905.02244, 2019.

[22] PRAJIT R, BARRET Z, QUOC V. Searching for activation functions [J]. arXiv Preprint, arXiv: 1710.05941, 2017.

[23] WOO S, PARK J, LEE J, et al. CBAM: convolutional block attention module [J]. arXiv Preprint, arXiv: 1807.06521, 2018.

[24] LI X, WANG W, WU L, et al. Generalized focal loss: learning qualified and distributed bounding boxes for dense object detection [J]. arXiv Preprint, arXiv: 2006.04388, 2020.

[25] HAMID R, NATHAN T, JUNY G, et al. Generalized intersection over union: a metric and a loss for bounding box regression [J]. arXiv Preprint, arXiv: 1902.09630, 2019.

[26] ZHENG Z, WANG P, LIU W, et al. Distance – IoU Loss: faster and better learning for bounding box regression [J]. arXiv Preprint, arXiv: 1911.08287, 2019.

[27] ZHENG Z, WANG P, REN D, et al. Enhancing geometric factors in model learning and inference for object detection and instance segmentation [J]. arXiv Preprint, arXiv: 2005.03572, 2020.

[28] ZHANG Y, REN W, ZHANG Z, et al. Focal and efficient IOU loss for accurate bounding box regression [J]. arXiv, Preprint arXiv: 2101.08158, 2022.

[29] CHATTOPADHAY A, SARKAR A, HOWLADER P, et al. Grad-cam ++: generalized gradient-based visual explanations for deep convolutional networks [C] // 2018 IEEE Winter Conference on Applications of Computer Vision (WACV). IEEE, 2018: 839–847.